

## Supplementary Material

### Principal Components Analysis

Although an analysis of variance (ANOVA) can be used to assess within-platform vs. between-platform variability on a gene-by-gene basis, it cannot be performed to analyze multiple genes simultaneously. For this purpose, we used a multivariate approach based on principal components analysis (PCA) (1) as visualized by Partek Pro (v. 5) software to summarize within-platform vs. between platform variability across a multidimensional scale. Like the comparison of the distribution of Z-scores, making comparisons between platforms with PCA requires that we bring all measurement technologies to a single common scale. In order to achieve this, we further applied a van der Waerden's transformation to the expression values from each chip for a direct comparison of platforms. This transformation entails replacing the data with its ranks and then applying an inverse normal transformation to the result to give (in the absence of tied values) a perfectly normal distribution. Supplemental Figure 3a shows the first two principal components of the data, which account for most of the variability in the multidimensional dataset ( $34.9 + 37.3 = 72.2\%$ ). Although the points (arrays) in supplemental figure 3a are color-coded by platform the principal components are computed without regard to platform category. It is clear from supplemental figure 3a that the between-platform variability is much greater than the within-platform variability (the first PC largely separating the cDNA from the two oligonucleotide platforms, and the second PC separating all platforms, especially the two oligonucleotide platforms). Supplemental figure 3a reveals that any variability due to the treatment condition is clearly overshadowed by the platform effect. Supplemental figure 3b shows a dot plot of PC #3, which clearly does capture the treatment effect. Since PC #3 explains 8.5% of the variability of the entire data, there does seem to be an ability to distinguish differential expression due to the treatment after masking out platform effects.

### Contingency Table Analysis

After modeling gene expression with an ANOVA model using a .001 alpha (Supplemental Table 6), we found that the Amersham assay detected the largest number of genes (117), the Agilent assay identified 67, and the Affymetrix assay found 34 differentially expressed genes. McNemar's test statistics computed for each of the three platform pairs indicated significant differences ( $p < .017$ ) in the number of genes found to be differentially expressed by each of the technologies. This observed difference may reflect the different levels of experimental variability associated with each of the platforms as seen in figure 3. For two of the comparisons, the Fisher's exact test for association between the gene sets were significant ( $p < .017$ ) demonstrating that agreement between the 2 platforms occurred in a manner that was non-random. However, when comparing the Amersham and Affymetrix gene lists, no significant (non-random) association could be found.

A subset of the genes in our lists exhibited statistically significant differential expression at an alpha cut-off of .001 with less than 2-fold differential expression. Since microarray technologies are frequently considered more accurate in detecting genes differentially expressed at levels greater than or equal to 2-fold (2), we applied an additional 2-fold change (in both directions) minimum criterion upon any genes found significant at a .001 alpha level to assess whether it would increase the level of overlap of differentially expressed genes detected by each platform. Despite applying this additional fold change criterion, we could not reject the null hypothesis ( $p < .017$ ) of no association between the Amersham and Affymetrix gene lists (Supplemental Table 7). Only when the alpha cutoff was reduced

to .01 with a 2-fold minimum criterion could we reject the null hypothesis of no association across all 2-way comparisons (Supplemental Table 8).

**Supplemental Table 1:** Pearson's product-moment and Spearman's rank-order correlation coefficients of gene expression measurements from 3 commercial microarray technologies matched by their Unigene ID. P-values of the hypothesis of no correlation are also reported.

Comparison	Platform A	Platform B	Pearson's	P-value	Spearman's	P-value	N
1	Amersham	Agilent	0.54505	<.0001	0.54630	<.0001	8024
2	Amersham	Affymetrix	0.59118	<.0001	0.58727	<.0001	8024
3	Agilent	Affymetrix	0.56198	<.0001	0.55232	<.0001	8024

**Supplemental Table 2:** Pearson's product-moment and Spearman's rank-order correlation coefficients of the fold change of time 0 hours and time 24 hours measurements matched by their Unigene ID. P-values of the hypothesis of no correlation are also reported.

Comparison	Platform A	Platform B	Pearson's	P-value	Spearman's	P-value	N
1	Amersham	Agilent	0.63374	<.0001	0.57991	<.0001	4012
2	Amersham	Affymetrix	0.56066	<.0001	0.53731	<.0001	4012
3	Agilent	Affymetrix	0.59903	<.0001	0.59549	<.0001	4012

**Supplemental Table 3:** Contingency tables of differential gene expression classifications from data matched by Unigene ID that was modeled with a mixed-model nested ANOVA and an alpha cut-off of .001. In each cell the frequency, row percentage, and column percentage is reported. Probability values from Fisher's Exact Test of association and McNemar's test of agreement are also reported.

Table of Agilent by Affymetrix			
Agilent	Affymetrix		Total
	No	Yes	
No	3800 98.52 96.50	57 <sup>+</sup> 1.48 77.03	3857
Yes	138 <sup>+</sup> 89.03 3.50	17* 10.97 22.97	155
Total	3938	74	4012

Table of Amersham by Affymetrix			
Amersham	Affymetrix		Total
	No	Yes	
No	3693 98.48 93.78	57 <sup>+</sup> 1.52 77.03	3750
Yes	245 <sup>+</sup> 93.51 6.22	17* 6.49 22.97	262
Total	3938	74	4012

Table of Amersham by Agilent			
Amersham	Agilent		Total
	No	Yes	
No	3650 97.33 94.63	100 <sup>+</sup> 2.67 64.52	3750
Yes	207 <sup>+</sup> 79.01 5.37	55* 20.99 35.48	262
Total	3857	155	4012

Platform A	Platform B	Exact P-value for McNemar's Test	Fisher's Exact Test P-value (Right)
Agilent	Affymetrix	<.0001	<.0001
Amersham	Affymetrix	<.0001	<.0001
Amersham	Agilent	<.0001	<.0001

\* significant non-random association

<sup>+</sup> significantly different detection rates

**Supplemental Table 4:** Pearson's product-moment correlation coefficients of technical and biological replicate measurements. P-values of the hypothesis of no correlation are also reported.

Comparison	Platform	Technical Replicates	P-value	Biological Replicates	P-value	N
1	Affymetrix	0.91894	<.0001	0.91255	<.0001	4018
2	Amersham	0.99259	<.0001	0.98240	<.0001	4018
3	Agilent	0.98727	<.0001	0.96435	<.0001	4018

**Supplemental Table 5:** Average difference of fold change measured by two platforms

Comparison	Platform A	Platform B	Mean Difference	N	Std Dev
1	Amersham	Agilent	0.0514661	2009	0.2502959
2	Agilent	Affymetrix	0.1033182	2009	0.3141808
3	Amersham	Affymetrix	0.1547843	2009	0.3347937

**Supplemental Table 6:** Contingency tables of differential gene expression classifications using a mixed-model nested ANOVA and an alpha cut-off of .001. In each cell the frequency, row percentage, and column percentage is reported, in that respective order. Probability values from Fisher's Exact Test of association and McNemar's test of agreement are also reported.

Table of Agilent by Affymetrix			
Agilent	Affymetrix		Total
	No	Yes	
No	1917 98.71 97.06	25 <sup>+</sup> 1.29 73.53	1942
Yes	58 <sup>+</sup> 86.57 2.94	9* 13.43 26.47	67
Total	1975	34	2009

Table of Amersham by Affymetrix			
Amersham	Affymetrix		Total
	No	Yes	
No	1863 98.47 94.33	29 <sup>+</sup> 1.53 85.29	1892
Yes	112 <sup>+</sup> 95.73 5.67	5 4.27 14.71	117
Total	1975	34	2009

Table of Amersham by Agilent			
Amersham	Agilent		Total
	No	Yes	
No	1848 97.67 95.16	44 <sup>+</sup> 2.33 65.67	1892
Yes	94 <sup>+</sup> 80.34 4.84	23* 19.66 34.33	117
Total	1942	67	2009

Platform A	Platform B	Exact P-value for McNemar's Test	Fisher's Exact Test P-value (Right)
Agilent	Affymetrix	0.0004	<.0001
Amerhsam	Affymetrix	<.0001	0.0441
Amersham	Agilent	<.0001	<.0001

\* significant non-random association

<sup>+</sup> significantly different detection rates

**Supplemental Table 7:** Contingency tables of differential gene expression classifications using a mixed-model nested ANOVA and an alpha cut-off of .001 and a 2-fold minimum criterion. In each cell the frequency, row percentage, and column percentage is reported, in that respective order. Probability values from Fisher's Exact Test of association and McNemar's test of agreement are also reported.

Table of Agilent by Affymetrix			
Agilent	Affymetrix		Total
	No	Yes	
No	1955 99.09 98.39	18 0.91 81.82	1973
Yes	32 88.89 1.61	4* 11.11 18.18	36
Total	1987	22	2009

Table of Amersham by Affymetrix			
Amersham	Affymetrix		Total
	No	Yes	
No	1933 98.98 97.28	20 <sup>+</sup> 1.02 90.91	1953
Yes	54 <sup>+</sup> 96.43 2.72	2 3.57 9.09	56
Total	1987	22	2009

Table of Amersham by Agilent			
Amersham	Agilent		Total
	No	Yes	
No	1933 98.98 97.97	20+ 1.02 55.56	1953
Yes	40 <sup>+</sup> 71.43 2.03	16* 28.57 44.44	56
Total	1973	36	2009

Platform A	Platform B	Exact P-value for McNemar's Test	Fisher's Exact Test P-value (Right)
Agilent	Affymetrix	0.0649	0.0005
Amersham	Affymetrix	<.0001	0.1236
Amersham	Agilent	0.0135	<.0001

\* significant non-random association

<sup>+</sup> significantly different detection rates

**Supplemental Table 8:** Contingency tables of differential gene expression classifications using a mixed-model nested ANOVA and an alpha cut-off of .01 and a 2-fold minimum criterion. In each cell the frequency, row percentage, and column percentage is reported, in that respective order. Probability values from Fisher's Exact Test of association and McNemar's test of agreement are also reported.

Table of Agilent by Affymetrix			
Agilent	Affymetrix		Total
	No	Yes	
No	1878 96.65 97.56	65 3.35 77.38	1943
Yes	47 71.21 2.44	19* 28.79 22.62	66
Total	1925	84	2009

Table of Amersham by Affymetrix			
Amersham	Affymetrix		Total
	No	Yes	
No	1794 96.66 93.19	62 <sup>+</sup> 3.34 73.81	1856
Yes	131 <sup>+</sup> 85.62 6.81	22* 14.38 26.19	153
Total	1925	84	2009

Table of Amersham by Agilent			
Amersham	Agilent		Total
	No	Yes	
No	1824 98.28 93.88	32 <sup>+</sup> 1.72 48.48	1856
Yes	119 <sup>+</sup> 77.78 6.12	34* 22.22 51.52	153
Total	1943	66	2009

Platform A	Platform B	Exact P-value for McNemar's Test	Fisher's Exact Test P-value (Right)
Agilent	Affymetrix	0.1078	<.0001
Amersham	Affymetrix	<.0001	<.0001
Amersham	Agilent	<.0001	<.0001

\* significant non-random association

<sup>+</sup> significantly different detection rates



**Supplemental Table 9:** Contingency tables of differential gene expression classifications using a mixed-model nested ANOVA and an alpha cut-off of .001. The Affymetrix data was normalized using dChip. In each cell the frequency, row percentage, and column percentage is reported. Probability values from Fisher's Exact Test of association and McNemar's test of agreement are also reported.

Table of Agilent by Affymetrix			
Agilent	Affymetrix		Total
	No	Yes	
No	1922 98.97 97.12	20 <sup>+</sup> 1.03 66.67	1942
Yes	57 <sup>+</sup> 85.07 2.88	10* 14.93 33.33	67
Total	1979	30	2009

Table of Amersham by Affymetrix			
Amersham	Affymetrix		Total
	No	Yes	
No	1875 99.10 94.74	17 <sup>+</sup> 0.90 56.67	1892
Yes	104 <sup>+</sup> 88.89 5.26	13* 11.11 43.33	117
Total	1979	30	2009

Platform A	Platform B	Exact P-value for McNemar's Test	Fisher's Exact Test P-value (Right)
Agilent	Affymetrix	<.0001	<.0001
Amersham	Affymetrix	<.0001	<.0001

\* significant non-random association

<sup>+</sup> significantly different detection rates

**Supplemental Table 10:** Contingency tables of differential gene expression classifications using a mixed-model nested ANOVA and an alpha cut-off of .001. The Affymetrix data was normalized using RMA. In each cell the frequency, row percentage, and column percentage is reported. Probability values from Fisher's Exact Test of association and McNemar's test of agreement are also reported.

Table of Agilent by Affymetrix			
Agilent	Affymetrix		Total
	No	Yes	
No	1914 98.56 97.11	28 <sup>+</sup> 1.44 73.68	1942
Yes	57 <sup>+</sup> 85.07 2.89	10* 14.93 26.32	67
Total	1971	38	2009

Table of Amersham by Affymetrix			
Amersham	Affymetrix		Total
	No	Yes	
No	1866 98.63 94.67	26 <sup>+</sup> 1.37 68.42	1892
Yes	105 <sup>+</sup> 89.74 5.33	12* 10.26 31.58	117
Total	1971	38	2009

Platform A	Platform B	Exact P-value for McNemar's Test	Fisher's Exact Test P-value (Right)
Agilent	Affymetrix	0.0022	<.0001
Amersham	Affymetrix	<.0001	<.0001

\* significant non-random association

<sup>+</sup> significantly different detection rates

**Supplemental Table 11:** Classifications based upon Affymetrix data normalized with dChip, RMA, and MAS5. Contingency tables of differential gene expression classifications using a mixed-model nested ANOVA and an alpha cut-off of .001. In each cell the frequency, row percentage, and column percentage is reported. Probability values from Fisher's Exact Test of association and McNemar's test of agreement are also reported.

Table of RMA by MAS5			
	MAS5		
	No	Yes	
No	1950 98.93 98.73	21 1.07 61.76	1971
Yes	25 65.79 1.27	13* 34.21 38.24	38
Total	1975	34	2009

Table of dChip by MAS5			
	MAS5		
	No	Yes	
dChip No	1950 98.53 98.73	29 1.47 85.29	1979
dChip Yes	25 83.33 1.27	5* 16.67 14.71	30
Total	1975	34	2009

Table of dChip by RMA			
dChip	RMA		Total
	No	Yes	
No	1952 98.64 99.04	27 1.36 71.05	1979
Yes	19 63.33 0.96	11* 36.67 28.95	30
Total	1971	38	2009

Platform A	Platform B	Exact P-value for McNemar's Test	Fisher's Exact Test P-value (Right)
RMA	MAS5	0.6587	<.0001
dChip	MAS5	0.6835	0.0001
dChip	RMA	0.3020	<.0001

\* significant non-random association

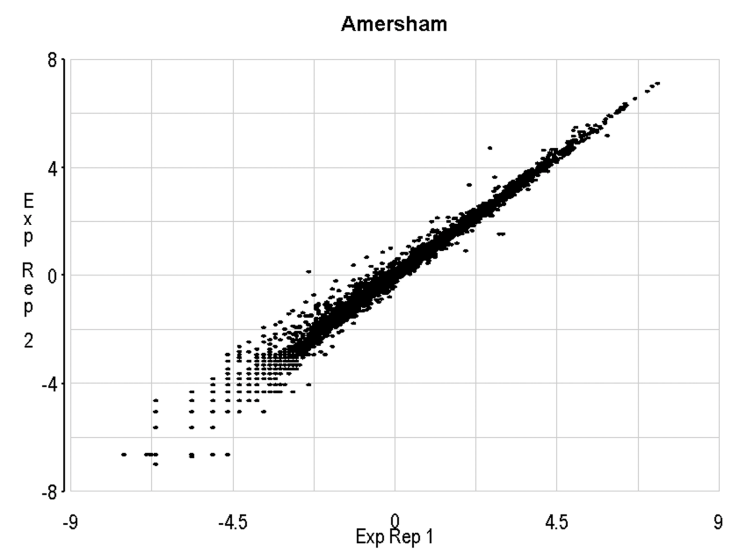
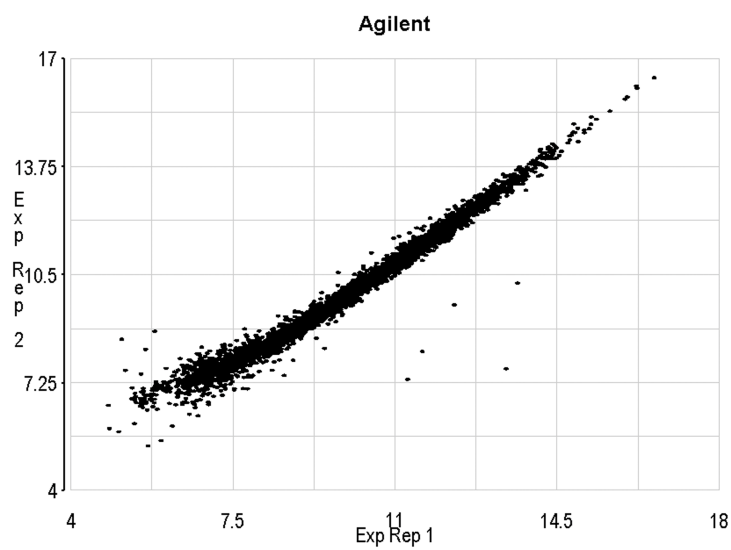
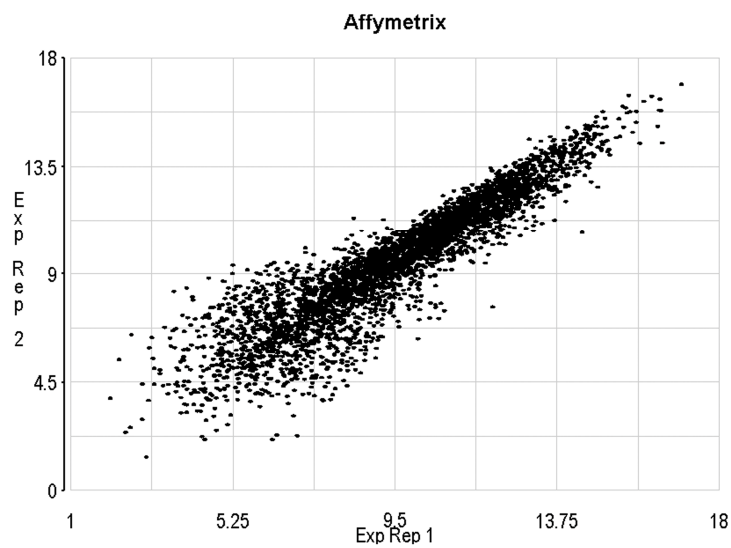
### **Supplemental Figure Legends**

**Supplemental Figure 1:** Scatter plots of log intensity values of the first and second experimental replicates.

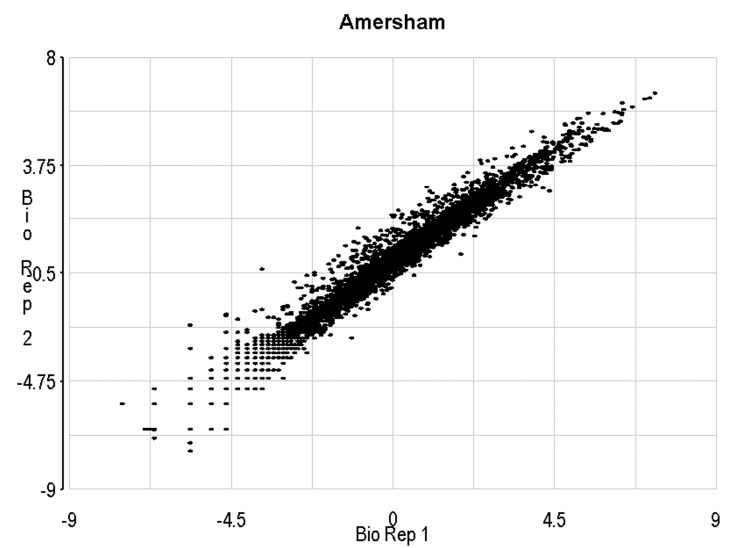
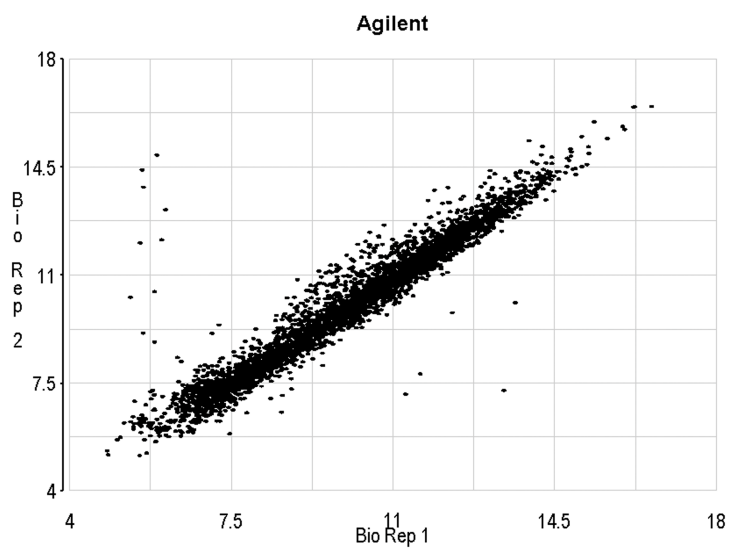
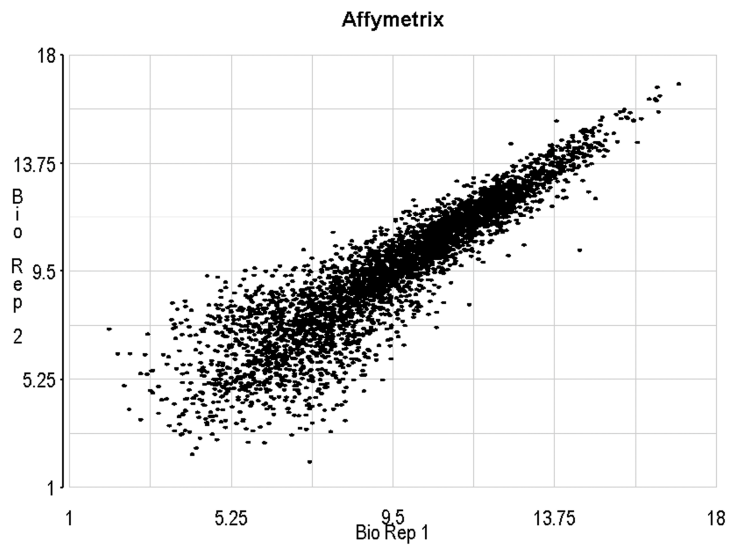
**Supplemental Figure 2:** Scatter plots of log intensity values of the first and second biological replicates.

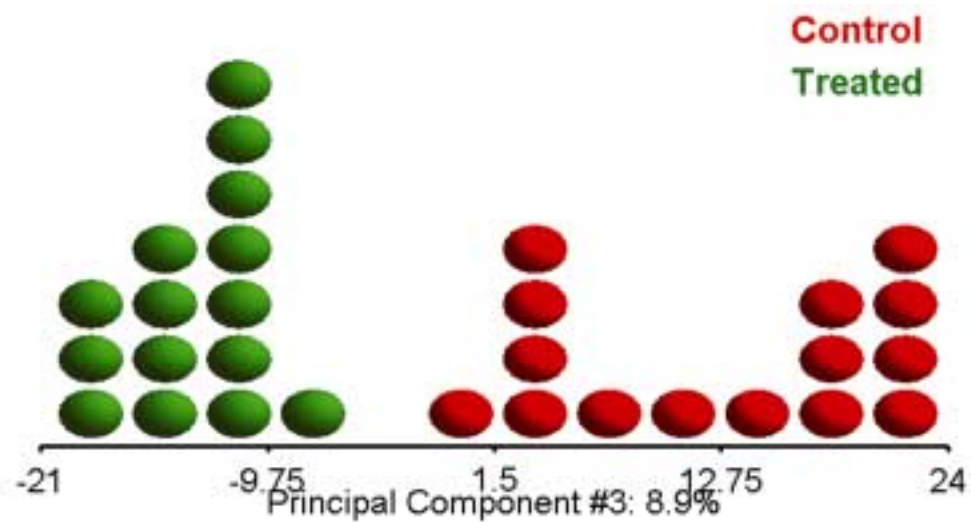
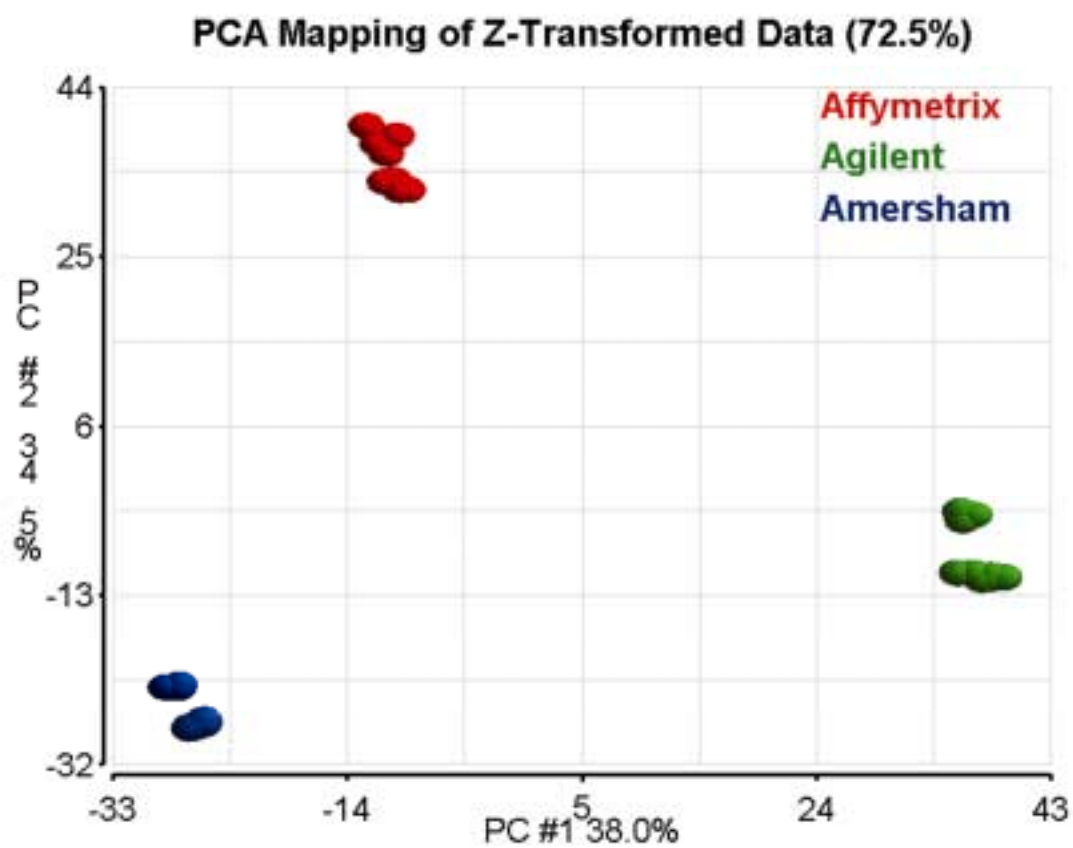
**Supplemental Figure 3:** Principal Components Analysis (PCA) of the data indicated that variation of signal values across microarray technologies was greater than signal variation caused by experimental treatment.

**Supplemental Figure 1**



**Supplemental Figure 2**





Supplemental Figure 3:

## References:

1. Jolliffe, I.T. (1986) *Principal Component Analysis*. Springer-Verlag, New York, NY.
2. Affymetrix Inc. (2002) GeneChip Expression Analysis Data Analysis Fundamentals.  
([http://www.affymetrix.com/Auth/support/downloads/manuals/data\\_analysis\\_fundamentals\\_manual.pdf](http://www.affymetrix.com/Auth/support/downloads/manuals/data_analysis_fundamentals_manual.pdf))